Development of Rheumatoid Arthritis

van de Stadt, L.A.

2013

document version
Publisher's PDF, also known as Version of record

13 A Prediction Rule for the Development of Arthritis in Seropositive Arthralgia Patients

L. A. van de Stadt [1]
B. I. Witte [2]
W. H. Bos [1]
D. van Schaardenburg [1,3]

[1] Jan van Breemen Research Institute | Reade, Amsterdam, The Netherlands
[2] VU University Medical Center, Department of Epidemiology and Biostatistics, Amsterdam, The Netherlands
[2] VU University Medical Center, Department of rheumatology, Amsterdam, The Netherlands

Ann Rheum Dis, 2012 nov 23, epub ahead of print

ABSTRACT

Objective

To predict the development of arthritis in anti-cyclic citrullinated peptide antibodies and/or IgM rheumatoid factor positive (seropositive) arthralgia patients.

Methods

A prediction rule was developed using a prospective cohort of 374 seropositive arthralgia patients, followed for the development of arthritis. The model was created with backward stepwise Cox regression with 18 variables.

Results

131 patients (35%) developed arthritis after a median of 12 months. The prediction model consisted of 9 variables: RA in a first degree family member, alcohol non-use, duration of symptoms <12 months, presence of intermittent symptoms, arthralgia in upper and lower extremities, VAS pain ≥50, presence of morning stiffness ≥1 hour, history of swollen joints as reported by the patient and antibody status. A simplified prediction rule was made ranging from 0 to 13 points. The AUC value (95%CI) of this prediction rule was 0.82 (0.75-0.89) after 5 years. Harrell's C (95% CI) was 0.78 (0.73-0.84). Patients could be categorized in three risk groups: low (0-4 points), intermediate (5-6 points) and high risk (7-13 points). With the low risk group as a reference, the intermediate risk group had a hazard ratio (HR; 95% CI) of 4.52 (2.42-8.77) and the high risk group had a HR of 14.86 (8.40-28.32).

Conclusions

In patients presenting with seropositive arthralgia, the risk of developing arthritis can be predicted. The prediction rule that was made in this patient group can help 1) to inform patients and 2) to select high-risk patients for intervention studies before clinical arthritis occurs.

INTRODUCTION

Rheumatoid Arthritis (RA) is an immune disease that can have detrimental consequences if left untreated. Apart from joint destruction with subsequent disability, RA patients also suffer from various comorbidities, resulting in increased mortality rate.[1,2] Joint destruction may already occur early in the disease course.[3] Early and aggressive treatment can be highly effective in controlling inflammatory activity and development of erosions[4-6] Some studies even suggest that treatment within this early 'window of opportunity' might alter the natural history of RA.[7,8] Early recognition of RA is thus highly important and several models have been created that predict the development of (erosive) RA in patients with undifferentiated arthritis.[9,10] Recently, the classification criteria for RA itself have been renewed to enable earlier classification.[11] However, these models and the new criteria concern patients that already have arthritis. To date no model exists to predict the development of RA before arthritis is clinically apparent.

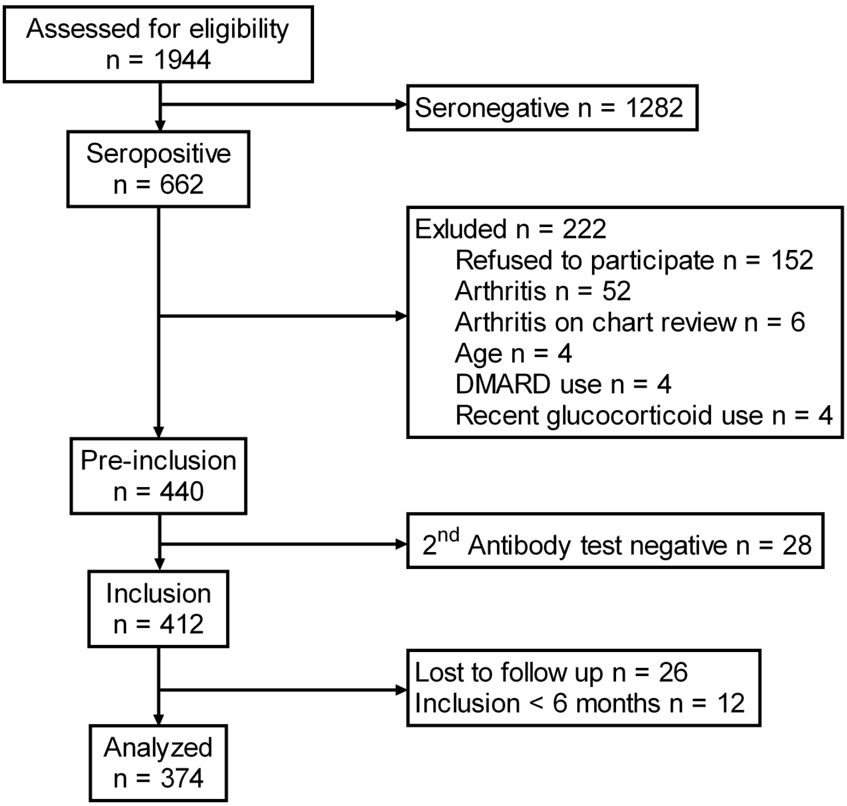
Patients presenting with arthralgia and a positive test for anti-cyclic citrullinated peptide antibodies (aCCP) and/or IgM rheumatoid factor (IgM-RF) (seropositive) are at risk for developing RA.[12] However, since not all of these patients develop arthritis[12] the question remains which of these patients will do so and within which time frame. We prospectively followed a cohort of seropositive arthralgia patients and made an easily applicable prediction rule consisting of nine clinical variables for the development of arthritis in these patients.

PATIENTS AND METHODS

Study population

Between August 2004 and June 2011, patients with a positive aCCP and/or IgM-RF status and (a history of) arthralgia, but not arthritis were recruited at rheumatology outpatient clinics in the Amsterdam area of the Netherlands.[12] Absence of arthritis was confirmed by physical examination of 44 joints by a trained medical doctor (WB or LAS) and a senior rheumatologist (DS).[13] Patients with arthritis as revealed by chart review or baseline physical examination, a negative aCCP and IgM-RF status on second analysis, age >70 years, previous treatment with a disease modifying antirheumatic drug (DMARD) or recent glucocorticoid treatment (<3 months) were excluded. Patients without follow-up due to loss or recent inclusion (follow-up less than 6 months) were excluded. **Figure 1** is a flowchart of inclusion. In total, 374 seropositive arthralgia patients were analyzed. Of these patients, 83 were also included in a randomized placebo-controlled trial studying the effects of two intramuscular dexamethasone injections on arthritis development. Since dexamethasone did not delay or prevent arthritis these patients were considered suitable for the present analysis.[14]

Figure 1. Flow chart of inclusion.



At baseline, medical history, details of joint complaints and the number of tender joints at physical examination were recorded.[15] Patients were seen semiannually the first year and annually thereafter. Extra visits were planned if arthritis developed. Development of arthritis in any of 44 joints was independently confirmed by two investigators (WB or LAS and DS).[13] Median follow-up was 32 months (inter quartile range (IQR): 13-48 months).

This study was approved by the local ethics committee and all participants gave informed consent.

Laboratory investigations

aCCP and IgM-RF levels were determined at baseline by second-generation aCCP ELISA (Axis Shield, Dundee, UK) and inhouse ELISA, respectively, as described previously.[16] The cut-off level for aCCP positivity was set at five arbitrary units/ml (AU/ml), according to the manufacturer's instructions. The cut-off level for IgM-RF positivity was set at 30 IU/ml determined on the basis of the analysis of receiver operating characteristic (ROC) curves.[16] C reactive protein (hsCRP) was measured using a highly sensitive latex-enhanced assay on a Hitachi 911 analyzer (Roche Diagnostics) and the cut-off level for CRP positivity was set at 10 mg/l, according to the manufacturer's instructions. HLA genotyping: HLA-DQ typing was performed as described previously.[17] HLA-DRB1 SE carrier status (one or two copies of the HLA-DRB1*0101, *0102, *0401, *0404, *0405, *0408, *0410 or *1001 alleles) was inferred from HLA-DQA1, -DQB1 haplotypes using strong linkage disequilibrium with HLA-DRB1 alleles in Caucasians.[18]

Statistical analysis

Data evaluation and statistical analysis were performed with SPSS V.20.0 software and the R software environment for statistical computing (R-Development Core Team). Prediction variables were selected based on clinical applicability, biologic plausibility, previous research and expert opinion. If variables were strongly correlated, one was selected to use in the prediction model, which resulted in 18 prediction variables (**table 1**). No variable had missing data exceeding 5%. Missing data were imputed with multiple imputation, creating 5 imputation sets. Baseline variables were assessed for their ability to predict remission by univariate and multivariate Cox proportional hazard analysis performed on the 5 imputation sets. Results of the pooled analysis are reported. Cases were treated as censored after the date of their last visit. Continuous variables with a non-linear association with the outcome variable were categorized using clinically applicable cut offs or percentiles. Categories were pooled if corresponding regression coefficients were similar. IgM-RF and aCCP status and titres were combined into one variable with four categories based on previous research (**Table 1**).[12] In the multivariate analysis all variables were entered in the model and subsequently selected with a backward stepwise procedure (p removal 0.1). Nagelkerke's R^2 was used to calculate the proportion of explained variation by this model.

A simplified prediction rule was obtained by rounding the regression coefficients to half points, which were then multiplied by 2 for easier clinical applicability. The diagnostic performance of the prediction model and rule were evaluated using the area under the curve (AUC) of receiver operating characteristic (ROC) curves and Harrell's C in the original dataset. ROC analysis was performed for arthritis at one, three and at five years. Patients with less than one, three or five years of follow-up, respectively, were treated as missing in the ROC analyses. Harrell's C is a performance measure that also takes into account time until development of arthritis, estimating the probability of concordance between the predicted and observed responses.[19] Values around 0.5 indicate no predictive discrimination, values close to 1, high predictive discrimination.

For internal validation, cross validation was performed, in which 300 randomly drawn cases from each imputation set were used as the test sets to create the prediction model with Cox regression, using the pooled analysis over the 5 datasets. The remaining 74 cases were subsequently used as the validation set to calculate the AUC of the ROC curves and Harrell's C as described above. This procedure was repeated 1000 times. The average scores from these 1000 analysis for the resulting regression coefficients, AUC and Harrell's C are presented.

RESULTS

Arthritis development

In total 374 patients, 76% female, mean ±SD age 49 ±11, were included in the present study. Patients had arthralgia for a median (IQR) of 12 (8-46) months with a median (IQR) number of reported painful joints of 4 (1-8). The median (IQR) tender joint count of 53 joints at physical examination was 0 (0-3). One hundred twenty patients (32%) were IgM-RF positive and aCCP negative, 143 patients (38%) were IgM-RF negative and aCCP positive and 111 patients (30%) were IgM-RF and aCCP positive. One hundred and ninety-four patients (52%) were shared epitope (SE) positive.

One hundred thirty-one patients (35%) developed arthritis after a median (IQR) follow-up of 12 (6-23) months. One hundred twenty-one (92%) of these patients could be classified as having RA according to the 2010 ACR/EULAR criteria, while 56 (45%) could be classified as having RA according to the 1987 criteria. Patients developing arthritis had a median (IQR) tender joint count of 5 (2-9) and a median (IQR) swollen joints count of 3 (2-5) at the time of diagnosis.

Table 1: Univariate analysis for patients that did develop arthritis and patients that did not*

Variable	Arthritis Yes (n = 131)	Arthritis No (n = 243)	HR (95% CI)	p
Age in years, mean (±SD)	47 (±11)	49 (±12)	0.99 (0.98 – 1.00)	0.20
Female sex, nr (%)	97 (74)	188 (77)	0.95 (0.64 – 1.42)	0.81
FDR with RA, nr (%)	38 (29)	48 (20)	1.26 (0.86 – 1.85)	0.23
Current smoker, nr (%)	46 (35)	67 (28)	1.11 (0.77 – 1.60)	0.57
No alcohol, nr (%)	58 (44)	78 (32)	1.60 (1.11 – 2.29)	0.01
NSAID use, nr (%)†	44 (34)	57 (24)	1.50 (1.04 – 2.57)	0.03
Duration of symptoms <12 months, nr (%)	48 (37)	72 (30)	1.47 (1.02 – 2.11)	0.04
Intermittent symptoms present,nr(%)	68 (52)	64 (26)	2.20 (1.56 – 3.11)	<0.001
Symmetric arthralgia, nr (%)	102 (78)	174 (72)	1.31 (0.87 – 1.98)	0.20
Arthralgia in upper and lower extremities, nr (%)	80 (60)	109 (45)	1.64 (1.15 – 2.34)	0.006
Arthralgia in small joints, nr (%)‡	96 (73)	164 (68)	1.07 (0.72 – 1.60)	0.76
VAS pain ≥ 50, nr (%)	59 (45)	69 (28)	1.87 (1.32 – 2.64)	<0.001
Morning stiffness ≥1 hour, nr (%)	31 (24)	37 (15)	1.94 (1.29 – 2.91)	0.001
Tender joint count, median (IQR)	0 (0–3)	0 (0–2)	1.06 (1.01 – 1.10)	0.005
Swollen joint(s) reported, nr (%)§	65 (52)	61(25)	2.42 (1.72 – 3.42)	<0.001
CRP positive, nr (%)	14 (11)	22 (9)	1.29 (0.73 – 2.26)	0.38
Antibody status				<0.001
IgM-RF positive, aCCP negative	12 (9)	108 (44)	1 (ref)	
aCCP low positive, IgM-RF negative¶	19 (15)	46 (19)	3.22 (1.56 – 6.64)	
aCCP high positive, IgM-RF negative||	33 (25)	45 (18)	4.73 (2.44 – 9.20)	
aCCP and IGM-RF positive	67 (51)	44 (18)	8.87 (4.79 – 16.4)	
Shared epitope positive, nr (%)	83 (63)	121 (50)	1.69 (1.17 – 2.42)	0.005

*HR: Hazard Ratio; CI: Confidence Interval; SD: standard deviation; nr: number; FDR: First Degree Relative; RA: rheumatoid arthritis; NSAID: non steroid anti-inflammatory drug; VAS: visual analog scale; IQR: interquartile range; CRP: C-reactive protein; IgM-RF: IgM rheumatoid factor; aCCP: anti-cyclic citrullinated peptide; † NSAID use: use of any NSAID in any dosing schedule; ‡ small joints: PIPs, MCPs, wrists and MTPs, § presence of a swollen joint as noticed by the patient, but not verified by a rheumatologist; ¶ aCCP low positive: ≥ 1x cut-off < 3x cut-off; || aCCP high positive ≥ 3x cut-off

Univariate analysis

To study which baseline characteristics could predict arthritis development, univariate Cox regression was performed first. Patients that developed arthritis versus those that did not develop arthritis were more often non-alcohol drinkers, used NSAIDs more often, more often had symptoms less than 12 months, intermittent symptoms, symptoms in both upper and lower extremities, a visual analogue scale (VAS) pain ≥50, morning stiffness ≥1 hour and reported swollen joints, had a higher mean tender joint count, and were more often aCCP and SE positive (**Table 1**).

Multivariate analysis

A prediction model for the development of arthritis was created with multivariate Cox regression using a backward stepwise approach (p removal 0.1). Variables included in the model and their regression coefficients are tabulated in **table 2**. The variables age, sex, smoking, NSAID use, symmetric symptoms, symptoms in small joints, tender joint count, CRP and Shared Epitope were excluded. When a lower p removal (0.05) was used, the variables first degree relative (FDR), alcohol non-use and location in upper and lower extremities were also excluded (data not shown). Performance measures of this model were slightly lower than of the full model (**Table 3**) and Nagelkerke's R^2 was 0.29.

Table 2. Multivariate analysis of predictors for arthritis development*

Variable	B	HR (95% CI)	B CV	HR (95% CI) CV	pts
FDR with RA	0.43	1.53 (1.04 – 2.26)	0.44	1.55 (1.24 – 1.96)	1
No alcohol	0.37	1.45 (0.97 – 2.17)	0.38	1.46 (1.20 – 1.79)	1
Duration of symptoms < 12 months	0.54	1.72 (1.18 – 2.52)	0.54	1.72 (1.39 – 2.15)	1
Intermittent symptoms present	0.55	1.74 (1.20 – 2.51)	0.56	1.75 (1.41 – 2.16)	1
Location in upper and lower extremities	0.32	1.37 (0.95 –1.99)	0.32	1.37 (1.12 – 1.68)	1
VAS pain ≥ 50	0.82	2.28 (1.57 – 3.30)	0.83	2.29 (1.85 – 2.85)	2
Morning stiffness ≥ 1hour present	0.50	1.65 (1.07 – 2.54)	0.51	1.67 (1.29 – 2.15)	1
Swollen joint(s) reported†	0.57	1.77 (1.24 – 2.53)	0.58	1.78 (1.46 – 2.20)	1
Antibody status					
IgM-RF positive, aCCP negative					0
aCCP low positive, IgM-RF negative‡	1.00	2.71 (1.31– 5.64)	1.00	2.71 (1.89 – 4.01)	2
aCCP high positive, IgM-RF negative§	1.58	4.85 (2.46 – 9.58)	1.60	4.95 (3.53 – 7.21)	3
aCCP and RF positive	1.94	6.94 (3.68 – 13.1)	1.96	7.09 (5.20 – 10.3)	4

*B: Regression coefficient; CV: cross validation; for other abbreviations see **table 1**;
† presence of a swollen joint as noticed by the patient, but not verified by a rheumatologist;
‡ aCCP low positive: ≥ 1x cut-off < 3x cut-off; §aCCP high positive ≥ 3x cut-off

For internal validation cross, validation was used. Regression coefficients did not change markedly after internal validation (**table 2; B CV**). As an alternative validation method, bootstrapping was also performed, using 1000 bootstrap samples to fit the prediction model. Regression coefficients did not change markedly either in these analyses (data not shown). The fraction of explained variation for the model (Nagelkerke's R^2) was 0.31.

To see whether the model changed if only aCCP positive patients were analysed, the model was also fitted to this patient group alone. This resulted in the same variables in the model with only slightly changed regression coefficients (data not shown).

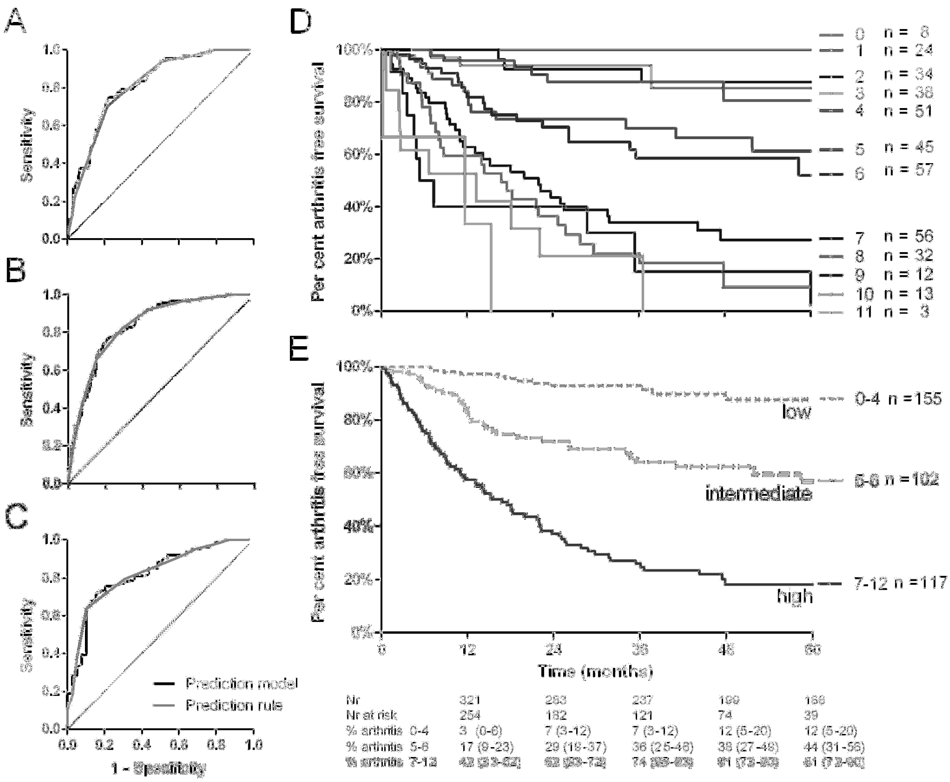


Figure 2. ROC and survival curves. A, B, C: The diagnostic performance of the prediction model (black line) and rule (grey line) were evaluated using the area under the curve (AUC) of receiver operating characteristic (ROC) curves. ROC analysis was performed for arthritis at one year (**panel A**), at three years (**panel B**) and at five years (**panel C**). Patients with less than one, three or five years of follow-up, respectively, were treated as missing in the ROC analyses.
D, E: The prediction rule resulted in a score for each patient. The arthritis free survival curves per risk score are shown in **panel D**. Patients could be categorized into three risk groups (panel E): 0-4 points: low (dotted line), 5-6 points: intermediate (dashed line) or 7-12 points: high risk (continuous line). Numbers at inclusion for each category are given on the right of **panel D** and **E**. Numbers in follow-up, numbers at risk and the percentage of arthritis development (95% CI) per time point are tabulated below figure 2E.

Prediction of arthritis.

For each patient a risk score was calculated by multiplying each variable with the regression coefficient of the prediction model. To create a more easily applicable prediction rule, regression coefficients were rounded to half points and multiplied by 2 (**table 2 and figure 3**). Multiplying each variable by the corresponding risk points resulted in a prediction rule score for each patient. Out of a potential score of 0 to 13 points, the scores in this cohort ranged from 0 to 12 points. The ROC curves for the prediction model and the rule for development of arthritis within one year (panel A), three years (panel B) and five years (panel C) are depicted in **figure 2**. The curves overlapped each other in every time interval, indicating that simplifying the model into the rule did not induce loss of discriminative ability. The AUC values and Harrell's C are tabulated in **table 3**.

Table 3. AUC values of ROC curves and Harrell's C*

	Harrell's C	AUC at 1 year	AUC at 3 years	AUC at 5 years
Prediction model	0.79 (0.74 – 0.84)	0.81 (0.76 – 0.87)	0.83 (0.78 – 0.88)	0.82 (0.74 – 0.89)
Cross validation	0.77 (0.68 – 0.85)	0.79 (0.67 – 0.89)	0.80 (0.67 – 0.90)	0.79 (0.60 – 0.94)
Prediction rule	0.78 (0.73 – 0.84)	0.81 (0.76 – 0.86)	0.83 (0.78 – 0.88)	0.82 (0.75 – 0.89)
Smaller prediction model†	0.77 (0.72 – 0.82)	0.80 (0.75 – 0.86)	0.81 (0.76 – 0.87)	0.79 (0.71 – 0.87)
Prediction model 2010 RA criteria	0.79 (0.74 – 0.85)	0.82 (0.77 – 0.87)	0.83 (0.78 – 0.89)	0.82 (0.75 – 0.90)
Prediction model 1987 RA criteria	0.77 (0.69 – 0.85)	0.80 (0.72 – 0.88)	0.81 (0.74 – 0.88)	0.81 (0.72 – 0.90)

* AUC: area under the curve; ROC: receiver operating characteristic. Values are presented with the 95% confidence intervals between brackets. †Prediction model created with p removal 0.05

Using Cox regression with the calculated risk score of the prediction rule as independent variable, the HR (95% CI) of the prediction rule per risk point was 1.58 (1.46-1.72). In **figure 2D**, the arthritis free survival curves are shown for all patients receiving a score from 1 to 11 – the one patient receiving 12 points was censored after 6 months and is not shown. As is clear from the survival curves, patients could be categorized into three risk groups: patients receiving 0 to 4 points having a low risk, patients receiving 5 to 6 points having intermediate risk and patients receiving 7 to 12 points having high risk. One hundred fifty-five patients (41%) had low risk, 102 patients (27%) had intermediate risk and 117 (31%) had high risk. The distribution of the variables with the highest weight (aCCP status and VAS ≥50) over the different risk categories is given in online supplementary **tables S1 and S2**. Hazard ratio's (95% CI) for the intermediate risk and the high risk group (with the low risk group as a reference) were 4.52 (2.42- 8.77) and 14.86

(8.40-28.32), respectively. **Figure 2E** depicts the arthritis free survival curves per risk group. After one year, 3% of the low risk group developed arthritis, 16% of the intermediate group developed arthritis, and 43% of the high risk group developed arthritis. After three years, these percentages were 7%, 36% and 74%, respectively. After five years these percentages were 12%, 43% and 81%, respectively (**Figure 2E**).

1.	Does the patient have a first degree relative with rheumatoid arthritis?	If yes	**1** point_____
2.	Does the patient drink alcohol?	If **no**	**1** point_____
3.	Did the symptoms start less than 12 months ago?	If yes	**1** point_____
4.	Are the symptoms intermittent?	If yes	**1** point_____
5.	Are there symptoms in both upper and lower extremities?	If yes	**1** point_____
6.	Is the VAS pain higher than or equal to 50?	If yes	**2** points____
7.	Does the patient have morning stiffness lasting more than or equal to one hour?	If yes	**1** point_____
8.	Did the patient notice any swelling in a joint?	If yes	**1** point_____
9.	Antibody status:		
	IgM-RF positive and aCCP negative		**0** points
	IgM-RF negative and aCCP positive < 3x cut-off		**2** points
	IgM-RF negative and aCCP positive >= 3x cut-off		**3** points
	IgM-RF and aCCP positive		**4** points____
			Total _____

Points	Risk	Arthritis development				
		1 year	2 years	3 years	4 years	5 years
0 to 4 points	Low	3%	7%	7%	12%	12%
5 to 6 points	Intermediate	17%	29%	36%	38%	44%
7 to 13 points	High	43%	63%	74%	81%	81%

Figure 3 Risk calculation form. With the risk calculation form, the total risk score can be calculated. Patients can be categorized into the low risk (0-4 points), intermediate risk (5-6 points) or high risk (7-13 points) group according to their total score.

RA classification criteria.

Most patients diagnosed with arthritis (92%), could be classified as having RA according to the 2010 ACR/EULAR criteria (RA2010). With RA2010 as the dependent variable in the Cox regression analysis, regression coefficients changed only slightly and the level of significance was a little lower for all variables (data not shown). This would have resulted in exclusion of the variable 'location in upper and lower extremities' because this variable had a p-value of 0.11. The explained fraction of variation of the prediction model (including the variable 'symptoms in upper and lower extremities') for RA2010 was 0.31 and with RA2010 as the dependent variable the AUC values of the ROC curves and Harrell's C did not change notably (**Table 3**).

Only 56 (45%) of the patients that developed arthritis could be classified as having RA according to the 1987 ACR criteria (RA1987). Cox regression analysis with RA1987 as the dependent variable would have resulted in a much smaller model, because the variables: FDR, alcohol non-use, location in upper and lower extremities and swollen joints reported by the patient all had p-values above 0.1. But since the number of events was lower, this could be due to loss of power. The explained fraction of variance of the full model for RA1987 was 0.19. The AUC and Harrell's C decreased only slightly.

DISCUSSION

Seropositive arthralgia patients are at risk of developing arthritis and subsequent RA. However, antibodies alone do not have sufficient predictive power to guide clinical decisions. We have developed a prediction rule for the development of arthritis in these patients, consisting of nine clinical variables that can easily be assessed during the first visit to the rheumatologist: RA in a first degree family member, alcohol non-use, duration of symptoms shorter than 12 months, presence of intermittent symptoms, arthralgia in both upper and lower extremities, VAS pain ≥50, presence of morning stiffness ≥1 hour, history of swollen joints as reported by the patient and antibody status. The discriminative ability was good with an AUC value of 0.82 and a Harrell's C of 0.79. Internal validation showed no major problems with overfitting. This rule can be readily used in clinical practice, since no complicated or expensive tests are necessary. It can help to inform patients and help to guide treatment decisions, although treatment protocols in these patients should first be tested in clinical trials.

For clinical applicability, a concern might be that in this study we chose the development of arthritis as outcome. However, arthritis can be self-limiting and still might not need DMARD treatment. We therefore also took the 2010 and 1987 RA criteria into account as outcome variables. The prediction rule then still had good discriminative abilities, although the proportion of explained variance of the prediction model for the 1987 RA criteria was low. An alternative outcome measurement would be erosive arthritis. However, in the present cohort, follow-up data during the arthritis phase on radiographic progression is still too little to provide reliable results. Therefore we did not take erosive arthritis into account as outcome variable, although it will be interesting to explore the discriminative ability of this prediction rule for erosive arthritis in future studies.

Another concern is that, while the prediction rule could discriminate high risk patients very well from low risk patients, there was still a group of patients with intermediate risk comprising 27% of the cohort. These patients had a risk of approximately 40% to develop arthritis within 5 years, which is considerable, but also still fairly uncertain. In such a patient group it might be useful to do additional tests such as genetics, RNA profiling, antibody profiling or other test methods that are currently being developed. Also, a more detailed study of the quality of the symptoms experienced by patients in this phase of the disease may be helpful. This has recently been identified as a research priority by the EULAR study group for risk factors for RA.[20] Additionally, different imaging techniques to assess the joints such as ultrasound, PET and MRI might prove to be useful in these patients.[21,22]

An interesting finding in this patient group was that alcohol consumption seems to protect against development of arthritis, as we and others have previously reported.[23-26] Although we would not advise patients to consume large amounts of alcohol, alcohol intake seems beneficial and, as our analysis indicates, should be taken into account in considering the risk of arthritis.

Factors that were surprisingly lacking in the model were presence of an acute phase response, presence of the SE, and smoking. We have previously shown that CRP levels and sPLA2 levels rise prior to development of RA in comparison to healthy individuals, but stay within the normal range.[27,28] In the present cohort, we have also observed that the CRP levels were slightly higher than in healthy blood donors but still within normal limits. Therefore, an elevated CRP level was not associated with arthritis development.[29] The presence of the SE on the other hand, was strongly associated with arthritis development in the univariate analysis. However, this association was not present in the multivariate analysis. An explanation for this phenomenon can be that the contribution of the SE to RA development is mediated via the presence and perhaps levels of ACPA.[30] The same could be true for smoking, which is highly associated with the formation of ACPA. The method of patient selection used for this study may cause the loss of association with arthritis of such risk factors within this cohort. We chose to develop a prediction rule for seropositive arthralgia patients from out of a clinical perspective and think the prediction rule will be a useful tool for this patient group. However, there are other groups of individuals at risk for development of arthritis who might be

seronegative. Therefore, it is essential that the prediction rule that we developed is validated in other prospectively followed cohorts of patients at risk for the development of RA, such as first degree family members or patients with inflammatory type arthralgia. Such cohorts are currently being set up and followed and it will be exciting to see the results of these studies and the comparison with our cohort.
In conclusion, we developed a prediction rule for the development of arthritis in seropositive arthralgia patient that is discriminative and easily applicable. It is a first step towards better understanding of and care for this patient group.

ACKNOWLEDGMENTS

The study was financially supported by the Dutch Arthritis Association, Grant number 0801034. We would like to gratefully thank all study participants for their efforts. Furthermore, we want to thank Margret de Koning and Roel Heijmans for their technical assistance.

SUPPLEMENTARY MATERIAL

Supplementary table S1: Distribution of variables aCCP status and VAS pain over the different risk categories.

Risk	CCP categories			VAS pain ≥ 50	
	negative	1–3x cutoff	≥3x cutoff	no	yes
Number	120	82	172	246	128
Low (0–4)	106 (88%)	28 (34%)	21 (12%)	12 (49%)	34 (27%)
Intermediate (5–6)	13 (11%)	29 (35%)	60 (35%)	73 (30%)	29 (23%)
High (7–10)	1 (1%)	25 (31%)	91 (53%)	52 (21%)	65 (51%

Supplementary table S2: Arthritis development in patients with high aCCP and a VAS ≥ 50

	Arthritis yes	Arthritis no
	131	243
High CCP and VAS ≥ 50	37(28%)	14 (6%)
Negative or Low CCP and/or VAS < 50	94 (72%)	229 (94%)

REFERENCES

1. Berkanovic,E. & Hurwicz,M.L. Rheumatoid arthritis and comorbidity. *J Rheumatol* 1990;**17:**888-892
2. Wolfe,F. *et al.* The mortality of rheumatoid arthritis. *Arthritis Rheum* 1994;**37:**481-494
3. Plant,M.J., Jones,P.W., Saklatvala,J., Ollier,W.E. & Dawes,P.T. Patterns of radiological progression in early rheumatoid arthritis: results of an 8 year prospective study. *J Rheumatol* 1998;**25:**417-426
4. Boers,M. *et al.* Randomised comparison of combined step-down prednisolone, methotrexate and sulphasalazine with sulphasalazine alone in early rheumatoid arthritis. *Lancet* 1997;**350:**309-318
5. Mottonen,T. *et al.* Comparison of combination therapy with single-drug therapy in early rheumatoid arthritis: a randomised trial. FIN-RACo trial group. *Lancet* 1999;**353:**1568-1573
6. Goekoop-Ruiterman,Y.P. *et al.* Comparison of treatment strategies in early rheumatoid arthritis: a randomized trial. *Ann Intern Med* 2007;**146:**406-415
7. Finckh,A., Liang,M.H., van Herckenrode,C.M. & de Pablo,P. Long-term impact of early treatment on radiographic progression in rheumatoid arthritis: A meta-analysis. *Arthritis Rheum* 2006;**55:**864-872
8. Lard,L.R. *et al.* Early versus delayed treatment in patients with recent-onset rheumatoid arthritis: comparison of two cohorts who received different treatment strategies. *Am J Med* 2001;**111:**446-451
9. Visser,H., le Cessie,S., Vos,K., Breedveld,F.C. & Hazes,J.M. How to diagnose rheumatoid arthritis early: a prediction model for persistent (erosive) arthritis. *Arthritis Rheum* 2002;**46:**357-365
10. van der Helm-van Mil,A.H. *et al.* A prediction rule for disease outcome in patients with recent-onset undifferentiated arthritis: how to guide individual treatment decisions. *Arthritis Rheum* 2007;**56:**433-440
11. Aletaha,D. *et al.* 2010 rheumatoid arthritis classification criteria: an American College of Rheumatology/European League Against Rheumatism collaborative initiative. *Ann Rheum Dis* 2010;**69:**1580-1588
12. Bos,W.H. *et al.* Arthritis development in arthralgia patients is strongly associated with anti-citrullinated protein antibody status: a prospective cohort study. *Ann. Rheum. Dis.* 2010;**69:**490-494
13. van der Heijde,D.M., van 't,H.M., van Riel,P.L. & van de Putte,L.B. Development of a disease activity score based on judgment in clinical practice by rheumatologists. *J Rheumatol.* 1993;**20:**579-581
14. Bos,W.H., Dijkmans,B.A., Boers,M., van de Stadt,R.J. & van Schaardenburg,D. Effect of dexamethasone on autoantibody levels and arthritis development in arthralgia patients: a randomized trial. *Ann. Rheum. Dis.* 2010;**69:**571-574
15. Ritchie,D.M. *et al.* Clinical studies with an articular index for the assessment of joint tenderness in patients with rheumatoid arthritis. *Q. J Med.* 1968;**37:**393-406
16. Nielen,M.M. *et al.* Specific autoantibodies precede the symptoms of rheumatoid arthritis: a study of serial measurements in blood donors. *Arthritis Rheum.* 2004;**50:**380-386
17. Crusius JBA. The immunogenetics of chronic inflammatory and autoimmune disease. . 2002. thesis/dissertation; Free University, Amsterdam, the Netherlands.
18. Knipper,A.J. *et al.* HLA-DRB1,3,4,5 and -DQB1 allele frequencies and HLA-DR/DQ linkage disequilibrium of 231 German caucasoid patients and their corresponding 821 potential unrelated stem cell transplants. *Hum. Immunol.* 2000;**61:**605-614
19. Harrell,F.E., Jr., Califf,R.M., Pryor,D.B., Lee,K.L. & Rosati,R.A. Evaluating the yield of medical tests. *JAMA* 1982;**247:**2543-2546
20. Gerlag,D.M. *et al.* EULAR recommendations for terminology and research in individuals at risk of rheumatoid arthritis: report from the Study Group for Risk Factors for Rheumatoid Arthritis. *Ann Rheum Dis* 2012;**71:**638-641
21. van de Stadt,L.A. *et al.* The value of ultrasonography in predicting arthritis in auto-antibody positive arthralgia patients: a prospective cohort study. *Arthritis Res Ther* 2010;**12:**R98

22. Gent,Y.Y. *et al.* Macrophage positron emission tomography imaging as a biomarker for preclinical rheumatoid arthritis: findings of a prospective pilot study. *Arthritis Rheum* 2012;**64:**62-66
23. van de Stadt,L.A. & van Schaardenburg,D. Alcohol consumption protects against arthritis development in seropositive arthralgia patients. *Ann Rheum Dis* 2012;**71**:1431-2
24. Maxwell,J.R., Gowers,I.R., Moore,D.J. & Wilson,A.G. Alcohol consumption is inversely associated with risk and severity of rheumatoid arthritis. *Rheumatology (Oxford)* 2010;**49:**2140-2146
25. Kallberg,H. *et al.* Alcohol consumption is associated with decreased risk of rheumatoid arthritis: results from two Scandinavian case-control studies. *Ann Rheum Dis* 2009;**68:**222-227
26. Pedersen,M. *et al.* Environmental risk factors differ between rheumatoid arthritis with and without auto-antibodies against cyclic citrullinated peptides. *Arthritis Res Ther* 2006;**8:**R133
27. Nielen,M.M. *et al.* Increased levels of C-reactive protein in serum from blood donors before the onset of rheumatoid arthritis. *Arthritis Rheum* 2004;**50:**2423-2427
28. Nielen,M.M. *et al.* Simultaneous development of acute phase response and autoantibodies in preclinical rheumatoid arthritis. *Ann Rheum Dis* 2006;**65:**535-537
29. Limper,M. *et al.* The Acute-phase Response Is Not Predictive for the Development of Arthritis in Seropositive Arthralgia - A Prospective Cohort Study. *J Rheumatol* 2012;39: 1914-7
30. van der Helm-van Mil,A.H. *et al.* The HLA-DRB1 shared epitope alleles are primarily a risk factor for anti-cyclic citrullinated peptide antibodies and are not an independent risk factor for development of rheumatoid arthritis. *Arthritis Rheum.* 2006;**54:**1117-1121